Position Statement: Penny Rheingans **September 10, 2004**

In the past few decades, the visualization community has developed elegant and effective techniques for the visual representation of many types of generally well-behaved data. It is now time for the field of visualization to move beyond simply adding to the collection of methods and to truly become a discipline with a principled approach to addressing the variety, complexity, ambiguity, and reality of real-world visualization needs and opportunities. Three key challenges before the visualization community are to gracefully incorporate the different forms of uncertainty that characterize data about an ambiguous world, to address computational and cognitive overload through the development and application of the semantics of visualization abstraction, and the objective evaluation of the nature and effectiveness of results.

Ambiguity. Data sampled from the real world is not unambiguous, because the world itself is fraught with ambiguity. Data samples themselves may be subject to uncertainty about spatial or temporal location, as well as value. Computational models which seek to explain influences and processes may produce exact numbers, but only if the complexity of variability has been smoothed away. Increasingly, computational models try to quantify both outcomes and the confidence in those predications. Visualization methods that ignore ambiguity cannot help but misrepresent the data they depict. We must move beyond special techniques for visualizing uncertainty to considering data quality aspects as a part of all visualization methods. Visualization frameworks that incorporate multiple sources and forms of uncertainty as a matter of course will produce visualizations that are more truthful and therefore more useful.

Abstraction. The abundant information available from modern sources is both a blessing and a curse: a blessing because more data potentially holds more clues to relationships and mechanisms and a curse because current data quantities threaten to outstrip our capabilities to explore and understand. Existing research has begun to address the computational challenges of truly huge datasets, but has only touched on conquering the cognitive overload from dense displays. Put more clearly, just because we can display a million items directly in a representation doesn't mean that we should. Creating a meaningful abstraction of a large dataset is if by no means a simple task, requiring not merely a reduction in volume but also a preservation of essential structure. The challenge of creating meaningful abstractions will require us to look beyond graphics and visualization knowledge, drawing in understanding of perceptual and cognitive processes from psychology research, techniques for effective visual communication from art and illustration, automatic discovery and summarization from machine learning, and domain specific knowledge from application experts. Representational and visualization abstraction of large and complex datasets is a necessary piece of addressing scale issues and conveying a core understanding in the face of potentially overwhelming detail.

Evaluation. In the early days of the visualization field, creating a novel representation of something which could previously not be represented visually was clearly a research contribution, but we've long since moved beyond those days. Now, it is not sufficient for a representation to be novel, but we must understand the characteristics of that representation in terms of both the quantitative and qualitative aspects of discovery. The quantitative aspects of effectiveness are beginning to be addressed as user studies of the accuracy of judgments become more common. The qualitative influences of visualization method on discovery are subtler, creating the potential for hidden biases of interpretation. We must more systematically as ourselves not just how well can someone make this judgment using this technique, but also what types of judgments do different types of techniques enable. As a community, it is imperative that we develop resources common resources to support the evaluation and comparison of visualization methods. At a minimum, these common resources include open source code, public data sets, and experimental protocols. Only thorough examination of the characteristics and effectiveness of visual representations can we claim that visualization is useful for supporting decisions with real consequences.